Library Resources & Technical Services. 1991. vol. 35, no. 1 p. 25-32.
ISSN: 0024-2527
Link to publisher website.
Link to journal website.

Backlog Management: Estimating Resources Needed to Eliminate Arrearages

Sally A. Rogers

Samples were drawn from two "historic" cataloging backlogs (one general and one consisting of Slavic language materials) at the Ohio State University. Collection development, cataloging, and preservation evaluations of the samples were used to estimate how many items were still wanted and to determine the resources that would be required to process those items. The results of these studies were used to determine that reduction of the general backlog could be accomplished as part of the regular workflow because only approximately one-fourth of the materials were still wanted. However, reduction of the Slavic backlog would require a special project due to a higher retention rate and the language expertise required for processing.

Library Literature covering 1986 - 1989 lists eleven articles on cataloging backlogs; only six are listed for the period 1955 - 1985. These figures suggest that backlogs again are becoming the major concern that they were in the early 1950s, when the *Journal of Cataloging and Classification* devoted an entire issue (Fall 1951) to the topic. The Library of Congress recently developed an "arrearage reduction management plan" to address its mind-boggling backlog of 38 million unprocessed items. Implementation of the plan is dependent upon the appropriation of funds by Congress.[1]

That backlogs exist in many libraries is no secret — surveys have shown that they are common (although the term "backlog" has been defined in various ways). Behrens and Smith surveyed 112 academic libraries in the United States holding more than 250,000 volumes and found that backlogs existed in 44 (85%) of the 52 responding libraries.[2] Agnew, Landram, and Richards conducted a survey of Association of Research Libraries (ARL) member libraries in 1984 and found that 68 (77%) of the 88 responding libraries had backlogs representing from 1% to 10% of their total collections. More than three-fourths of the libraries with backlogs had had them over ten years.[3] These results were consistent with those obtained by Piternick when he surveyed ARL libraries and selected others in 1968. At that time, 67 of the 86 responding libraries (78%) had backlogs.[4] These studies imply that backlogs did not disappear during the 1960s and 1970s even though they were mentioned infrequently in the library literature.

A distinction must be made between "normal" and "historic" processing backlogs. "Normal" backlogs routinely develop whenever the volume of incoming materials is high and sometimes are used to stabilize the cataloging workflow when the volume is low. They also might be created purposefully by libraries preferring to wait for cataloging copy or authority records to become available from the bibliographic networks. "Normal" backlogs probably are desirable as long as they can be contained: that is, as long as periods of backlog growth are offset regularly by periods of backlog reduction.

In contrast, the "historic" backlogs found in many libraries often consist of thousands of volumes that have been awaiting cataloging for years. Of the 60 libraries that participated in the 1989 National Shelflist Count, 27 (45%) reported that they had processing backlogs of more than

20,000 volumes. The median backlog size for these 27 libraries was 70,000 volumes.[5] Backlogs of this size take up valuable storage space that possibly could be put to better use. Many of the materials are likely to be outdated and might no longer be worthwhile additions to the library's collections. Some volumes are unusable due to physical deterioration. While "historic" backlogs can stabilize the cataloging workflow in times of severe financial exigency when materials budgets are drastically reduced, they are not desirable. They are evidence of a prolonged imbalance in the scheme of technical processing that should be addressed by library administrators.

Backlogs develop when more materials are acquired than can be processed. There are various reasons why this imbalance occurs. White and Roos cite materials budgets, staffing levels in cataloging departments, special projects such as retrospective conversion, collection-development policies, and cataloging priorities of the Library of Congress as factors that can contribute to the growth of backlogs.[6] Another factor suggested by Dwyer is the quality of the records used for copy cataloging; when records must be corrected or improved before they can be used, cataloging is more time consuming.[7] Intner attributes the current backlog crisis to the increasing variety of materials being published, the conversion to computerized systems, a more complicated cataloging code, and shrinking processing budgets.[8]

Various methods of eliminating, controlling, and providing access to backlogs have been described by Halverson.[9] Harcourt,[10] Miller and Ford,[11] Moomaw, [12] and Share[13]. The reported strategies include attacking the backlog through special projects, integrating backlogged materials with the regular workflow, accepting more cataloging copy with less editing, providing temporary or abbreviated cataloging, expanding the role of paraprofessional staff, and improving the efficiency of workflows and procedures. These strategies are based on the assumption that everything in the backlog should be cataloged. If library administrators question whether materials acquired many years earlier are still worth adding to the collection, a "reselection" project like the one initiated by the Library of Congress in 1983 to weed out unneeded items can significantly reduce the number of materials requiring cataloging.[14]

When backlogs are large, old, and of uncertain origin, it can be difficult to know what approach to use and where to start to gain control. Intner suggests "bibliographic triage" might be necessary. She discusses several possible answers to the question of what should be cataloged first and stresses the importance of knowing the costs and long-range effects associated with each alternative.[15] Sampling can be a useful tool for assessing the nature of a large backlog in order to determine the cost and impact of various backlog management and elimination strategies. White and Roos describe a backlog assessment project conducted at the University of Wyoming Library in 1986. A random sample of 660 titles was surveyed using 16 variables to gather data to be used to improve backlog management techniques.[16]

This paper describes a backlog assessment project recently conducted at the Ohio State University Libraries (OSUL) to facilitate the elimination of two historic backlogs. The desire to provide improved access to these materials and the need to prevent them from physically deteriorating were two concerns that prompted the study. Uncataloged materials at OSUL are assigned control numbers and are included in the Libraries" online catalog, the Library Control System (LCS). While library patrons generally do not have physical access to these materials, author and title access is provided by LCS. Rush cataloging is done in response to any patron request for an unprocessed item (about 1,200 a year).

According to the survey conducted by Behrens and Smith, 63.6% of the academic libraries reporting that they had backlogs were providing some type of public access to their uncataloged

materials.[17] However, as Piternick wrote, "'Materials acquired by a library but whose presence in the library is inadequately recorded and displayed are of little use to the library's users.[18] Even when brief record access to backlog materials is provided, the real issue is still adequacy of access. Generally, backlogs consist of library materials that are less accessible than they should be. The technological advances that are rapidly improving access to all kinds of information are raising the expectations of library users and making the continued existence of large historic backlogs unacceptable.

The Research Questions

A logical first step in coping with the challenges presented by historic backlogs is to estimate the resources that would be required to eliminate the arrearages. At OSUL a project was designed to evaluate random samples from two large monographic backlogs to which no materials had been added for over a decade. One backlog consisted of Slavic language materials; the other consisted primarily of social science and humanities materials in English and western European languages. Both the Slavic and general backlogs included fiction and nonfiction works. The goal of the project was to determine: (1) how many items were still wanted for the Libraries' collections; (2) how much cataloging copy was available for those items; and (3) how many were in poor physical condition. The results of these evaluations would help OSUL staff estimate the resources required to eliminate the two backlogs.

Data Collection and Analysis

Generating the Samples

Materials in the Slavic language backlog being studied all had control numbers beginning with "PRES," while those in the general backlog had numbers beginning with "PREA." Computer-generated random samples were drawn from LCS using these control number prefixes to define the populations to be studied. Although the sample sizes were not selected to ensure certain levels of confidence and precision, the samples were large enough to obtain a standard error of less than 6% with a 95% confidence level for all but the last stage of the study. For the final stage, which was conducted with substantially reduced sample sizes, a standard error of less than 6.5% with a 90% confidence level was obtained.

In early 1987, two samples of 300 titles were drawn from the records for the Slavic backlog (23,540 items). Since the brief LCS record used for uncataloged materials includes the publication date, it was possible to generate samples consisting of 300 titles published in or before 1945 and 300 published after 1945. This stratification by imprint date was recommended by a library faculty member with specialized knowledge of Slavic publications. An analysis of the publication dates of the Slavic backlog materials indicated that about 71% of the 23,540 items were post - 1945 imprints. Therefore, in taking a sample of 300 titles from each stratum, a proportionately larger sample of the pre - 1945 imprints was taken.

In early 1988, a sample consisting of 990 titles was taken from the LCS records for the general backlog materials (20,719 items). An analysis of the publication dates of the general

backlog materials indicated that about 25% of the items were post - 1945 imprints. No stratification by imprint date was recommended in sampling the general backlog.

Preliminary Searching

The control number, author, title, and publication date for each sample item were reported on a 3-by-5-inch card. The cards for each sample were arranged in control number order (the order in which the backlogged materials were stored. Student assistants stamped the cards with consecutive numbers beginning with one; the same numbers were stamped on the worksheets that would be used to record the results of bibliographic searching activities. The cards were matched to and inserted in the sample items as they were retrieved from the backlog storage area.

When library staff pulled the samples from storage, 23 (8%) of the pre - 1945 Slavic items, 16 (5%) of the post - 1945 Slavic items, and 93 (9%) of the general items could not be found. When these snags" were searched by author and title in LCS, it was discovered that half of the missing Slavic materials and almost a fourth of the missing general materials already had been cataloged. In these cases, the brief order-level record had been erroneously retained rather than properly replaced when the catalog record was added to LCS.

Sample items successfully retrieved from the storage area were searched in LCS to identify duplicates, i.e., additional copies of cataloged titles. If there was insufficient bibliographic information in LCS to confirm potential duplicates, the cataloged titles were retrieved from the stacks and compared to the sample items. (When other uncataloged copies of sample items were identified, they were retrieved from the storage area and t hereafter considered part of the sample.) One item (0.3%) from each stratum of the Slavic sample and 52 items (5%) from the general sample proved to be duplicates of previously cataloged titles. The duplicates were not evaluated further as a part of this study but were handled in a separate workflow.

Search for Cataloging Copy

The remaining sample titles (i.e., those that had not been identified as snags, duplicates, or as having been already cataloged) were searched in the Online Computer Library Center's (OCLC) Online Union Catalog to locate matching or related cataloging copy. No further searching was undertaken once "exact match" OCLC copy was found. Remaining titles were then searched in *the National Union Catalog (NUC).* The *Slavic Cyrillic Union Catalog (SCUC)* was also checked for pre - 1956 Slavic imprints if no copy was found in *NUC.* Titles for which exact match copy still was not found were searched in the Research Libraries Information Network (RLIN) database. The results of these searching activities are reported in table 1. Only 34% of the general items required searching beyond OCLC, compared to 77% of the pre - 1945 and 58% of the post - 1945 Slavic items. Almost all of the matching *NUC* records for both backlogs were found in the volumes for pre - 1956 imprints. Searching the subsequent volumes of *NUC* was found not to be worth the considerable amount of time it required. Similarly, searching RLIN proved not to be worthwhile since very little additional copy was found.

Due to the nature of the cataloging workflow at OSUL, searchers were instructed to indicate whether or not each matching record they found included a call number based on the *Library of Congress Classification* (copy lacking an LC-type call number must be sent to the

section responsible for original cataloging). Sixty percent of the matching copy for the pre - 1945 Slavic materials and 46% of the copy for the post - 1945 Slavic and the general materials included an LC-type call number.

Searchers also noted the existence of any related cataloging records (e.g., for a different edition) whenever they were unable to find exact match copy. Monographs for which matching or closely related copy is found are cataloged by the Copy Cataloging Section at OSUL. In some cases, a call number or subject headings must be assigned by the section responsible for original cataloging before processing can be completed. Monographs for which no copy is found may be given minimal level cataloging depending on the length of time they have been in the in-process area, the language and content of the text, and public service needs.

TABLE 1

RESULTS OF SEARCH FOR CATALOGING COPY

	Slavic Samples				General Sample	
	PRE - 1945		Post - 1945			
	no.	%	no.	%	no.	%
Exact match		(95%c.i.)		(95% c.i.)		(95% c.i.)
OCLC	63	22.8 ±4.96	119	42.0 ± 5.74	556	65.8 + 3.19
NUC	49	17.8 ±4,51	107	37.8 ± 5.64	173	20.5 j 2.72
SCUC	48	17.4 ±4.47	3	1.1 ±1.22	0	0.0
RLIN	11	4.0 ±2.31	5	1.8 ±1.55	6	0.7 ± 0.55
Subtotal	171	62.0 ±5.72	234	82.7 ±4.41	735	87.0 ± 2.27'
Related copy	32	11.6 ±3.78	12	4.2 ± 2.33	62	7.3 ±1.74
No copy	73	26.4 ±5.19	37	1.3.1 ±3.92	48	5.7 ±1.57
Total	276	100.0	283	100.0	845	100.0

Collection Development Evaluation

The sample titles from the Slavic backlog were evaluated by the East European and Slavic Studies Librarian; however, because the general backlog included a wide variety of materials, seven collection managers were invited to review the sample titles and select any they wanted to have added to the Libraries' collections. The project worksheets indicating the results of the searching done to identify duplicates and to locate cataloging copy were made available to the collection managers to assist them in making their decisions. Of the 276 pre - 1945 Slavic titles that were evaluated, 194 (70.3%) were selected for addition to the OSUL collections; 166 (58.7%) of the 283 post - 1945 Slavic titles evaluated were selected. Of the 845 general backlog titles that were evaluated, only 231 (27.3%) were selected. Ninety-five percent confidence intervals constructed around these sample proportions indicated that the true proportions of backlogged items that would be selected for retention would be between 64.9% and 75.7% of pre - 1945 Slavic titles, between 53% and 64.4% of post - 1945 Slavic titles, and between 24.3% and 30.3% of general titles.

Preservation Evaluation

The physical condition of each of the sample items selected for retention by the collection managers was evaluated by a library staff member assisting the preservation officer. A separate worksheet was used to record the results of this evaluation, which are shown in table 2. Each volume was checked for environmental damage (e.g., mold), defacement, brittle and / or acidic paper, missing pages, and weak or broken binding. A number of factors were considered in determining what type of preservation action should be taken when a volume was found to be in very poor condition. For example, replacement was recommended if a reproduction or reprint was available and the volume was expected to be used regularly. If demand was expected to be low and the volume included high quality illustrations that could not easily be reproduced, a protective box was recommended. Not surprisingly, about 76% of the pre - 1945 Slavic items selected for retention will require some kind of preservation action (e.g., binding, enclosure, or replacement). About 57% of the post - 1945 Slavic titles and half of the general titles will require such action.

TABLE 2

RESULTS OF PRESERVATION EVALUATION

	Slavic Samples				General Sample	
	Pre - 1945		Post - 1945			
	no.	% (90%c.i.)	no.	% (90%c.i.)	no.	% (90%c.i.)
No action needed	47	24.2 ± 5.08	72	43.4 ± 6.35	115	49.8 ± 5.43
Commercial binding	21	10.8 ± 3.68	62	37.4 ± 6.19	18	7.8 ± 2.90
Pamphlet binding	11	5.7 ± 2.74	19	11.4 ±4.08	12	5.2 ± 2.41
Repair in-house	10	5.2 ± 2.62	2	1.2 ±1.39	16	6.9 ± 2.76
Provide enclosure in-house	47	24.2 ± 5.08	7	4.2 ± 2.57	4	1.7 ±1.40
Replace or reformat	53	27.3 ± 5.28	4	2.4 ±1.96	55	23.8 ± 4.62
Other (e.g., clean)	5	2.6 ± 1.88	0	0.0	11	4.8 ± 2.33
Total	194	100.0	166	100.0	231	100.0

Data Applications

Based on the results of the analysis of the backlog samples, steps have been taken at OSUL to facilitate the elimination of the two historic backlogs. The collection development officer has asked certain collection managers to review a small number of items from the general backlog each week until the entire collection has been evaluated. Since less than 25% of these items are expected to be selected, no searching to identify duplicates or cataloging copy has been done in advance. Instead, the collection managers have been asked to search each title they select in the OSUL online catalog, making identification of duplicates part of their review process. The selected items will then be funneled into the regular cataloging workflow, which includes searching for cataloging copy and special procedures for handling materials in poor physical condition. The brief records in the Libraries' online catalog for the materials that are not selected for retention will be deleted, and the items will be donated (without further review) to the Friends of the Libraries for possible inclusion in the booksales they regularly sponsor.

The results of the analysis of the Slavic samples were shared with four faculty members from the Department of Slavic and East European Languages and Literatures who agreed to assist in evaluating and prioritizing the Slavic backlog. Initially, the faculty members will review a printout of the backlog titles listed in the Libraries" online catalog. They may also examine the corresponding physical pieces as necessary. Given the expected retention rate for the Slavic materials and the need for the selected items to be processed by staff with language expertise, some additional cataloging personnel will be required to achieve backlog reduction. Library administrators are negotiating with the department faculty members for some assistance to support the processing of the items they identify as high priority. Depending on the uniqueness and value of the materials selected from the Slavic backlog, it might also be possible to obtain outside funding to support processing of the collection.

Conclusion

For OSUL, sample assessment was an effective means of gathering information about two sizeable backlogs that had resulted from an extended imbalance between acquisitions and cataloging production. The backlog study provided the data needed to choose appropriate courses of action and to proceed with confidence that the Libraries' resources would be used to the best advantage. The results also underscored the need for backlog growth to be contained on an ongoing basis.

Analysis of the general sample showed that only approximately one-fourth of the general backlog was still needed and that cataloging copy was available for 87% of those titles. Clearly, if collection managers could evaluate the general backlog, a large amount of storage space could be made available fairly quickly; and the selected items could be handled as part of the regular workflow.

The analysis of the Slavic samples showed that a different approach would be needed to reduce the Slavic backlog. A much larger proportion of these materials would be selected for retention. This finding suggests that the Slavic titles might have been acquired more selectively than the general titles, at least some of which were known to have been acquired through large block purchases. The analysis of the Slavic samples also indicated that less cataloging copy would be available, particularly for the older materials, and that processing staff with language expertise would be required.

The Chi Square method of hypothesis testing was used to compare the results obtained by the cataloging copy, collection development, and preservation evaluations of each stratum with the results obtained for each of the other two strata in turn. It was possible to conclude with 95% confidence that the differences among the strata were statistically significant in each case.

Backlogs are a challenge faced by many libraries. To determine the best way to handle this challenge, library managers need to know what is in the backlog, how much of it is still wanted, and what it will take to add those items to the collection. Analysis of samples is one technique that can be used to provide this information.

Analyzing a sample can help to insure that the right people are involved in the effort to eliminate a backlog and that available resources are concentrated where they will do the most good. The data produced can indicate whether it is realistic to try to incorporate backlog reduction into the regular workflow or whether a special project will be required. The estimated

cost of eliminating a backlog can be computed using cost and production figures associated with the regular workflow. Backlog reduction efforts can then be prioritized along with other projects competing for budget allocations. If external funding must be sought, the sample data can be used to support grant proposals. Library managers who use a methodology like the one described above to analyze a backlog sample should be one step closer to meeting the backlog challenge.

References and Notes

1. "Library Develops Plan to Cut Arrearages," *Library of Congress Information Bulletin* 49:170, 174 (May 7, 1990).
2. Beth Behrens and Philip M. Smith, "Cataloging Backlogs in Academic Libraries," *Tennessee Librarian* 39:15 (Winter 1987).
3. Grace Agnew, Christina Landram, and Jane Richards, "Monograph Arrearages in Research Libraries," *Library Resources & Technical Services* 29:345 (Oct. 1985).
4. George Piternick, "University Library Arrearages," *Library Resources & Technical Services* 13:104-5 (Winter 1969).
5. *National Shelflist Count: Titles Classified by Library of Congress and National Library of Medicine Classifications, 1989.* (Chicago: Association for Library-Collections & Technical Services, American Library Assn., 1990).
6. Carol White and Tedine Roos, "Sampling the Cataloging Backlog: The University of Wyoming Library's Experience," *Technical Services Quarterly* 6, no. 1:16 (1988).
7. Jim Dwyer, "Harrumph! Wherein a Colonial Cataloger Speaks Out Against Backlogs and the British MARC Menace," *Technicalities* 6, no.5:1 (May 1986).
8. Sheila S. Intner, "Bibliographic Triage," *Technicalities* 7, no. 12:10 (Dec. 1987).
9. Jacque Halverson, Ava Nell Harris, and Leila Payne, "Bonus Babies: a Special Project." *Technicalities* 5, no. 11:12-14 (Nov. 1985).
10. Jill Harcourt, "Cataloging Rare Book Backlogs," *College & Research Libraries Mews* 47:127-29 (Feb. 1986).
11. Ruby Miller and Barbara J. Ford. " 'Ask at Reference' for Baeklogged Books," *College h Research Libraries Mens* 49:12-15 (Jan. 1988).
12. Judith Mooniaw. "The Light at the End of the Tunnel," *Journal of Educational Media it Library Sciences* 21:180-89 (Winter 1984).
13. Donald Share, "Management of Backlogs." *Library journal* 111, no. 14:160-61 (Sept. 1. 1986).
14. "Semiannual Report on Developments at the Library of Congress: October 1, 1982, through March 31, 1983," *Library of Congress Information Bulletin* 42:173 (May 30,1983). For additional reports on LC's reselection project, see the "Semiannual Report on Developments at the Library of Congress." *Library of Congress Information Bulletin* 42:437 (Dec. 19, 1983; 43:420 (Dec. 24, 1984); 44:153 (July 1, 1985); 45:5 (Jan. 6, 1986); and 45:215 (June 16, 1986).
15. Sheila S. Intner, "Bibliographic Triage-Revisited, " *Technicalities* 8, no. 10:3-4 (Oct. 1988).
16. White and Roos, "Sampling the Cataloging Backlog," p.17-18.
17. Behrens and Smith, "Cataloging Backlogs in Academic Libraries," p. 16.
18. Piternick, "University Library Arrearages," p. 103.